



HOME > 분석 > 과학

바이러스 확산 시뮬레이션 해보니...한국은 상대적 '안전지대'

지윤성 팩트체커 | 승인 2020.02.03 13:10 | 댓글 0



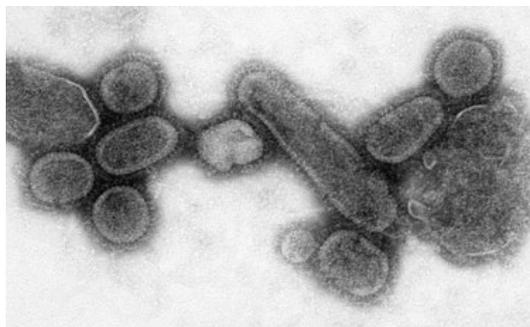
전염병 확산 시뮬레이션 프로그램인 STEM과 GLEAM 활용
 한국 최대 1만명 감염 가능성...5~6월이 누적 정점
 감염자 수는 중국-태국-인도네시아-베트남 순 예측
 GLEAM 모델의 경우 무증상 감염 가능성 인자도 고려 요소로 반영

신종 코로나바이러스 감염증 확산으로 전세계가 비상이 걸렸다. 전 세계 연구진은 빅데이터 및 머신러닝 기술을 활용하여 국지적인 전염병이 어떻게 인접국가로 확산되는지 시뮬레이션을 하고 있다. 뉴스톱은 시뮬레이션 작업이 어떻게 이뤄지는지 소개하고자 한다. 다만 시뮬레이션 결과는 제한된 정보를 활용한 예측이기 때문에 의사결정을 위한 보조수단 이상으로 활용되어서는 안된다는 점을 미리 알린다.

1. 바이오인포매틱스와 전염병 예측 모델의 발전 (출처)

신종 코로나바이러스 감염증이 확산되면서 각국이 선제적 대응에 나서고 있다. 이런 선제적 조치에는 여러 판단 근거들이 필요한데 최근에는 감염병 확산 예측 시뮬레이션 결과가 의사결정에 활용되고 있다. 이를 바이오인포매틱스(Bioinformatics)라고 한다.

바이오인포매틱스를 활용해서 질병 확산을 예측하는 기술은 2009년 3월 처음 알려진 H1N1(돼지독감) 전지구적 대유행(Pandemic)으로 촉발되었다. 당시 세계보건기구는 2009년 6월 11일 H1N1의 경보단계를 대유행을 의미하는 6단계로 선언하였다. 세계보건기구의 인플루엔자 대유행에 대한 선언은 1968년 이후 41년 만의 일이었다. 세계보건기구의 자료에 따르면 확진자 163만2258명이고, 감염에 의해 사망한 사람은 20여개국에서 나타났다. 감염환자가 나타난 나라는 129여개국으로 총 1만9633명이 사망한 것으로 집계 되었다. 물론 이 수치 역시 타임플루에 내성을 가진 변종까지 포함하고 역학조사를 통해 합병증 환자들까지 고려한다면 더 높을 수 밖에 없다.



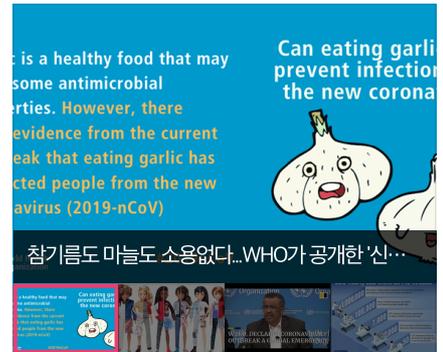
인플루엔자 A 바이러스 서브타입 H1N1 현미경 사진.

전 지구적인 대유행을 겪었던 2009년 H1N1 신종플루와 관련된 다양한 확산 데이터들과 결과들을 기본 학습데이터로 활용하여 예측 시뮬레이션 모델을 만들기 시작한 것이 미국의 질병통제예방센터(CDC)다. 초기에는 미국내 확산을 그래프 형식으로 나타낼 수 있는 시뮬레이션 프로그램 개발로 이어졌다. 대표적 사례인 '커뮤니티

최신뉴스

- 참기름도 마늘도 소용없다..WHO가 공개한 '신종 코로나'...
- 트랜스젠더 입학 반대? 생물학적으로 십여개의 성(Sex)이...
- [팩트체크] 신종 코로나보다 미국 독감이 더 심각하다?
- "미국 독감으로 신종코로나 물타기?"
- 민주-한국당 대표 맞고발.. '정치의 사법화'와 '사법의 정치'...

포토뉴스



인기뉴스

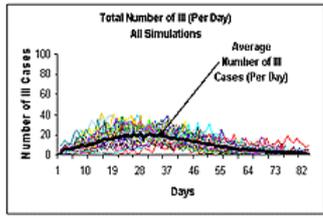
- 1 신종 코로나 바이러스의 '9가지 진실'
- 2 바이러스 확산 시뮬레이션 해보니...한국은 상대적 '안전지대'
- 3 1931년 우리인의 중국인 혐오, '130여명 학살'로 이어지
- 4 "미국 독감으로 신종코로나 물타기?"
- 5 [주간팩트체크] 중국인 입국 금지하면 코로나바이러스 막



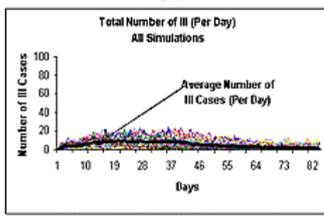
수를 입력하면 감염병 확진 환자수를 예측해 그래프 형식으로 제공한다. 하지만 이런 초창기 예측 프로그램은 실무자들이 직접 예측에 필요한 모든 변수들을 수기로 입력해야 했다. 간단한 수준의 계산만이 가능해 질병이 발생한 지역의 인구학적·지리학적 특성을 제대로 반영하기에는 무리가 있다.

Example of graphic output:

Cases without intervention:



Cases with intervention (School closings):



Community Flu 2.0 분석 결과 예시. 출처: CDC

7 [팩트체크] 신종 코로나보다 '미국 독감'이 더 심각하다?

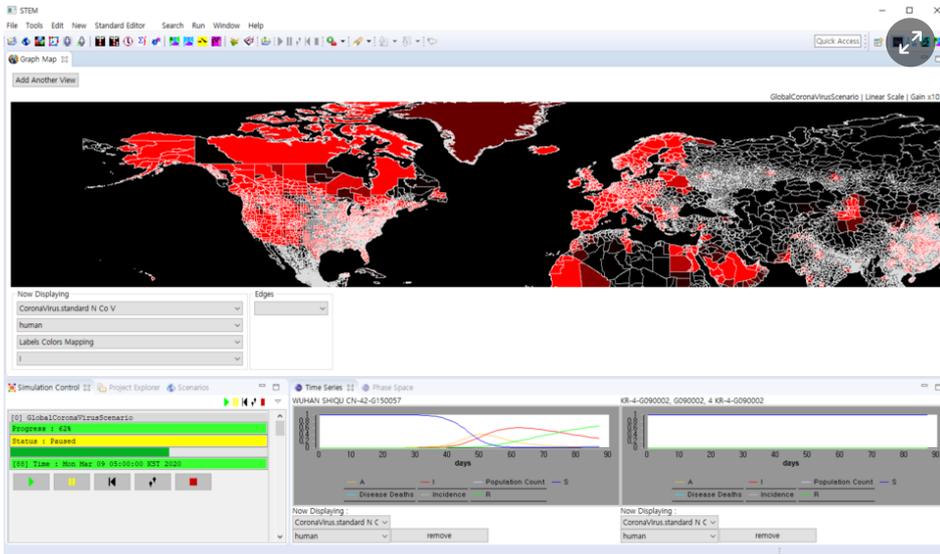


8 이 와중에 '반중 혐오' 부추겨 조희수 올리는 언론들



2. STEM과 GLEAM, 전염병 예측 모델 프로그램의 쌍두마차

초기 모델 이후 지리 정보와 인구집단의 상호작용 정보, 통근 인원 기록, 항공기 노선 정보를 접목해서 현실 세계를 좀더 잘 반영할 수 있는 모델이 나오기 시작했다. 인공지능과 빅데이터를 활용한 모델이다. 대표적인 모델 중 하나는 IBM이 개발하고 Eclipse재단에서 무료 배포중인 The Spatiotemporal Epidemiological Modeler(시공간 감염 확산 모델러, 일명 STEM)이다. 다른 하나는 유럽연합에서 주로 활용하는 Global Epidemic and Mobility project (글로벌 감염 확산 이동 모델 프로젝트, 일명 GLEAM)이다. 참고로 두 재단 모두 분석 소프트웨어를 해당 사이트에서 무료 배포하고 있고 다양한 연구 논문들에 활용되고 있는데 STEM의 경우 난이도가 상당히 높고 복잡한 반면 GLEAM의 경우 좀더 직관적이고 편하다.

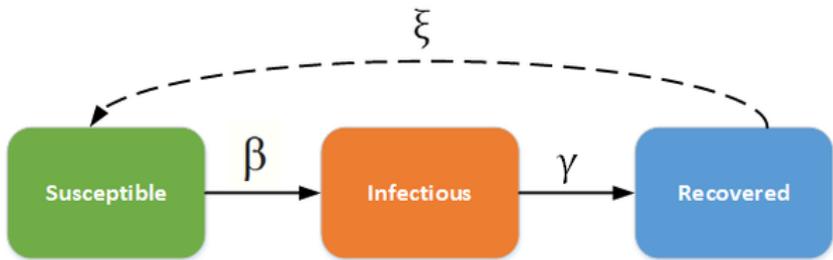


시공간 감염 확산 모델인 STEM 화면 캡처.



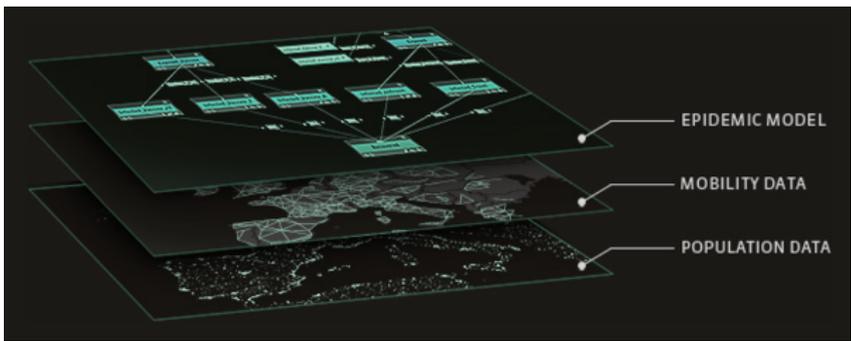
감염병을 예측하는 기술은 △수학적인 알고리즘 중심의 예측기술 △ 슈퍼컴퓨터와 같은 고성능 컴퓨팅 인프라 중심의 예측기술 △예측 대상 지역의 이동인구 패턴을 설명하는 빅데이터 연계 시뮬레이션 기반 기술 등의 단계를 거치며 진화하고 있다. 수식 기반의 이론역학의 기원은 1760년 네덜란드의 물리학자이자 수학자였던 다니엘 베르누이가 내놓은 천연두 발생과 관련한 수확모델에서 찾을 수 있다. 이후 1927년 영국의 커맥과 맥캔드릭이 제안한 질병구획 중심의 미분방정식은 지금까지 다양한 질병 확산 모델링에 사용되고 있다. 감염 확산 모델을 간략히 표현하면 아래와 같다.

$$D_k^n = \sum_{k'} D_{k'}^{n-1} (k' - 1) P(k|k') \left(1 - R_0^{-\lambda_{k'k}}\right) \left(1 - \sum_{m=0}^{n-1} \frac{D_k^m}{V_k}\right)$$

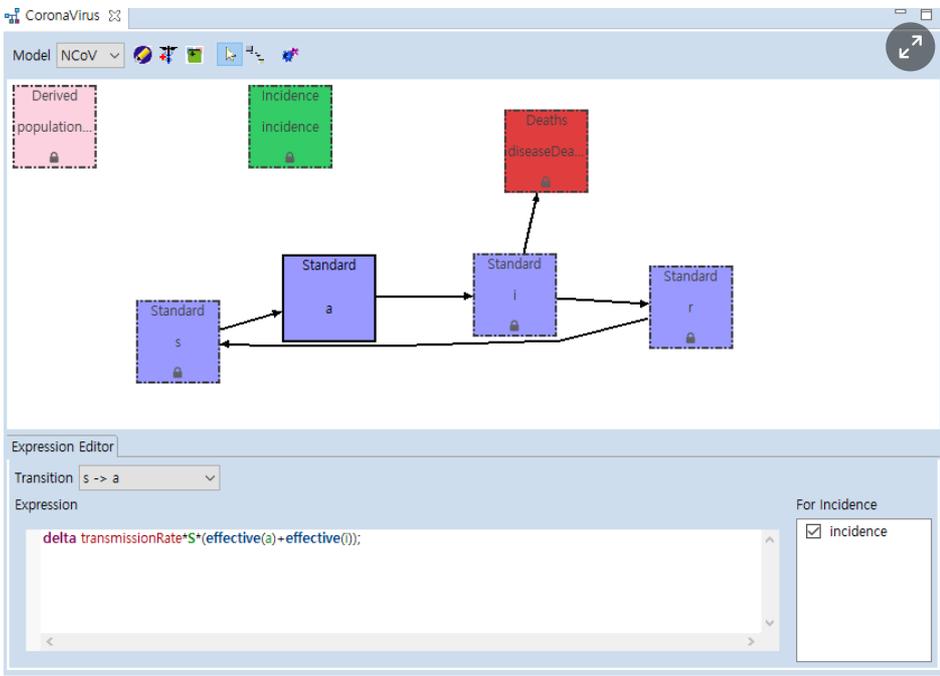


SIR 감염확산 모델

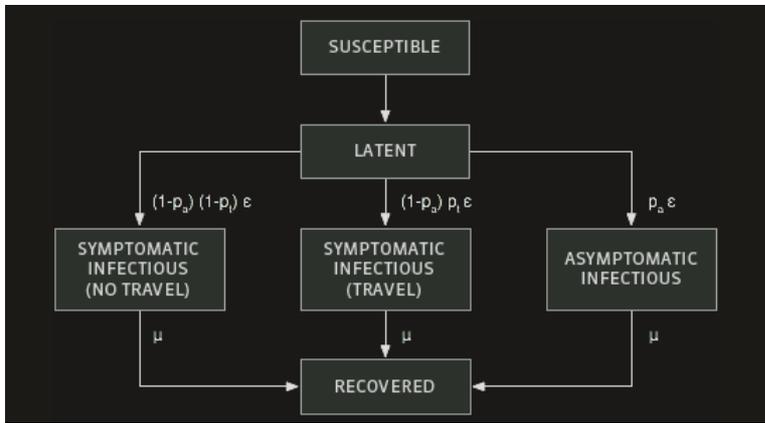
SIR 모델은 Susceptible(감염가능한 취약 계층)->Infectious(감염)->Recovered(회복) 단계를 의미한다. 이런 수식을 기본으로 하되 최근에는 다양한 통계 빅데이터들을 레이어링해서 분석한다. 인구 데이터 위에 이동 데이터, 그리고 전염병 모델을 더해서 확산 정도와 경로를 파악하는 것이다.



해당 모델을 기반으로 분석 시뮬레이션을 할 때 다양한 변수가 있다. 감염자도 똑같은 감염자가 아니라 Latent(잠복), Symptom-T(증상발현, 이동), Symptom-NT(증상발현, 이동없음), Asymptom(비증상) 등으로 분류할 수 있다. 분석 및 연구자의 노하우와 컴퓨팅파워에 따라 천차 만별의 결과가 나온다. 뉴스톱은 STEM과 GLEAM을 활용해 신종 코로나 바이러스 확산을 직접 시뮬레이션 해보았다. 적용한 모델은 CDC/DDPHSS의 브라이언 박사의 제안 모델을 이용하였고 GLEAM의 경우 H1N1 Pandemic 모델을 사용하였다.



STEM 적용 프레임워크



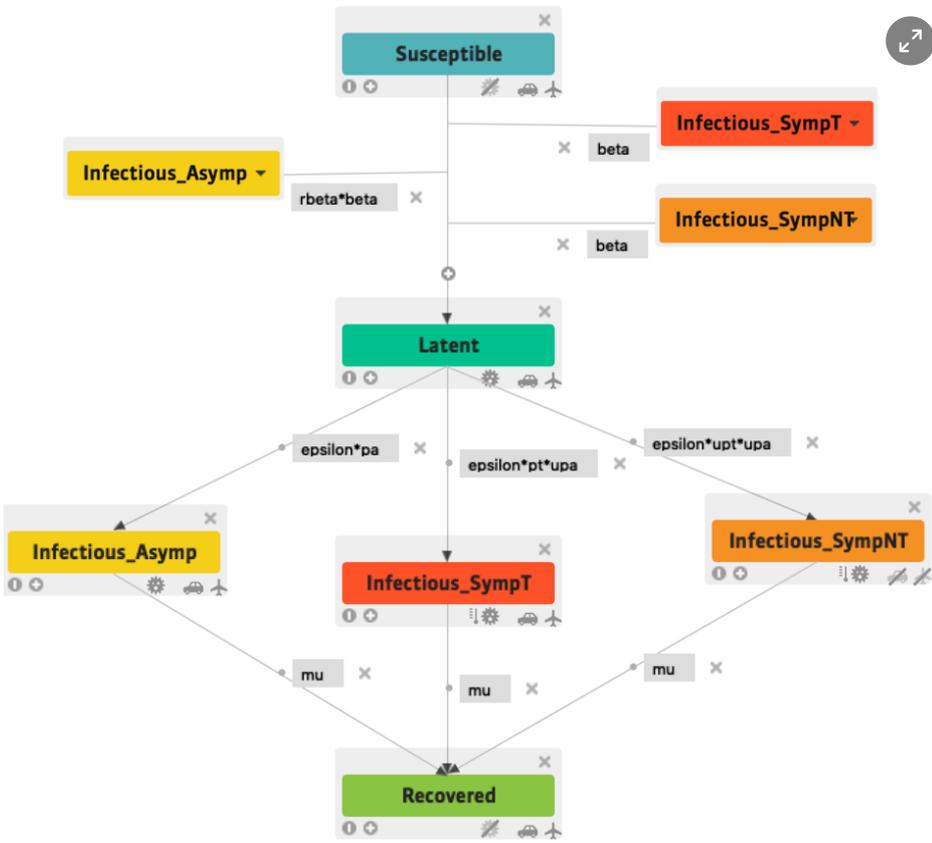
GLEAM 적용 프레임워크

3. 시뮬레이션 결과 중국 우한 신규 감염자 확산은 향후 1-2개월이 정점, 그외 국가는 6월 이후까지 확산 예상

이번 신종 코로나 바이러스 확산은 중국 정부가 대부분 정보를 가지고 있기 때문에 정확한 시뮬레이션에 한계가 있다. 최초 몇명이 어디서 감염됐는지 알 수 없고 중국이 내놓은 자료에 대한 신빙성도 의문시되고 있다. 뉴스톱이 보여주는 시뮬레이션 자료는 참고용으로만 사용해야 한다. 이 모델에서는 어느 지역으로 확산될 수 있는지, 감염증은 얼마나 오래 지속될지를 보여준다. 시뮬레이션은 동일한 조건을 가지고 STEM, GLEAM을 따로 따로 수행했으며 비교 해보았다.

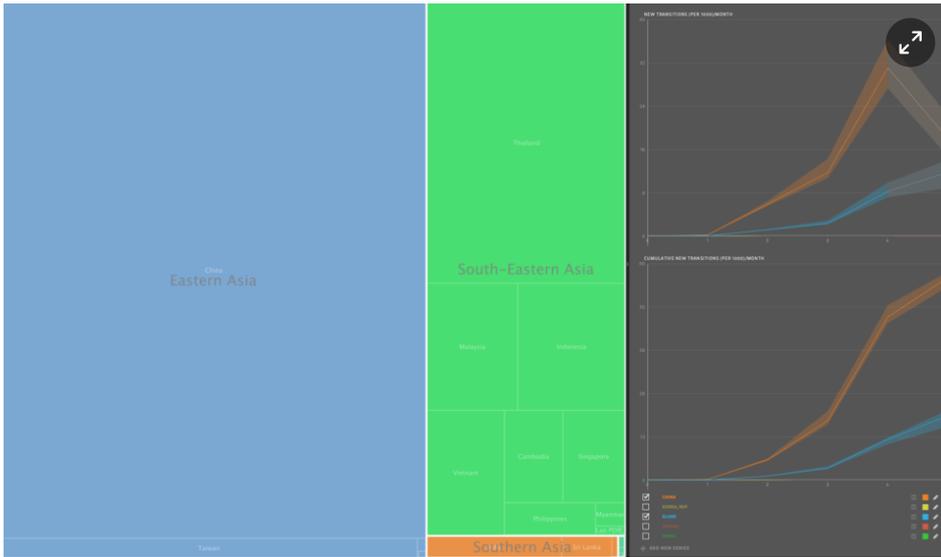
본 시뮬레이션에는 2월 2일 현재 존스홉킨스 대학의 신종 코로나 감염자수 통계와 함께 항공 이동, 지역내 통근 통계, 각 지역의 인구통계, 과거 신종 플루 및 코로나 바이러스 확산 자료, 이동 통제(Intervention)와 같은 교란 요소, 각국의 의료보건위험성지수와 정부위기대응지수 등 다양한 기관의 데이터를 망라하였으며 수학적으로는 룬게-쿠타(Runge-Kutta method)를 근사해(approximate solution) 찾기로 사용했다. 특히 UN산하 190여 개국의 국가별 보건인프라, 정부기구 및 대응 조직 취약성 등 자료를 취합해 변수로 넣었다. 세계 각국이 기존 전염성 질병에 어떻게 대응했는지도 확산 예측에 반영이 됐다는 의미다.

우선 유럽에서 주로 사용하며 비교적 다루기가 쉬운 GLEAM의 시뮬레이션 결과이다. 아직 이번 신종 코로나에 대한 각종 요인들이 밝혀지지 않아 기본 모델은 2009년 H1N1의 글로벌 확산모델을 사용하였으며 각종 계수들은 지금까지 알려진 신종 코로나 바이러스 감염 정보들을 활용하여 일부 수정하여 분석하였다.



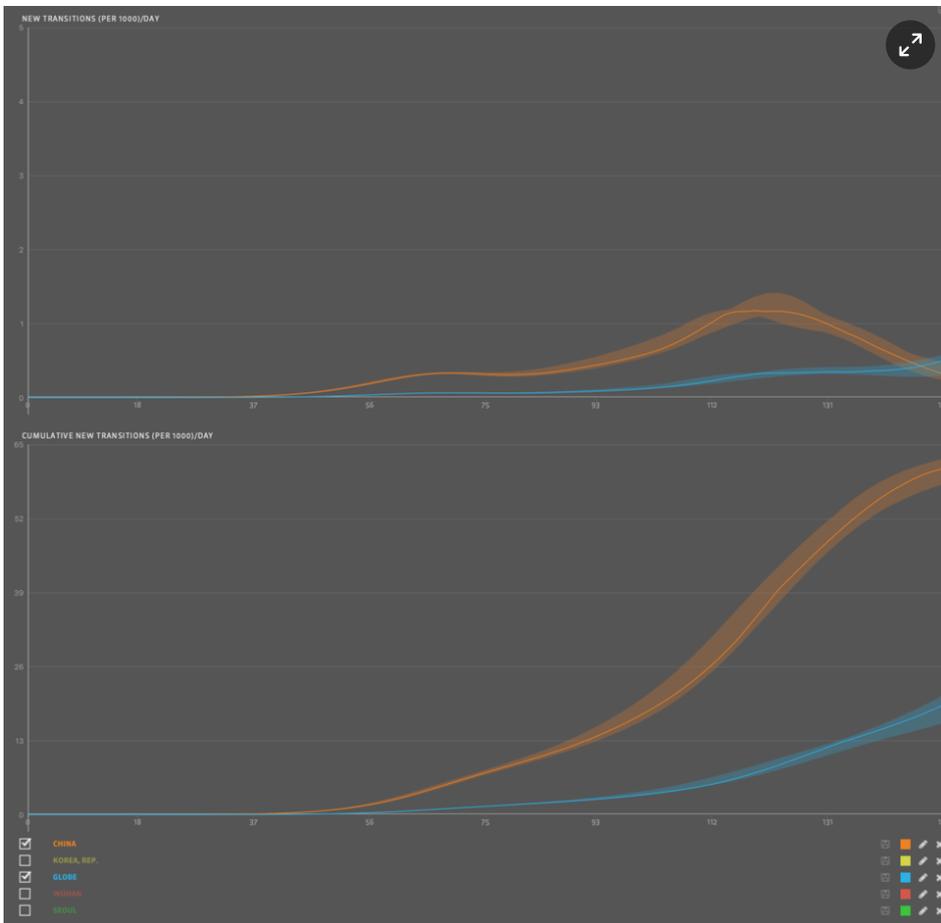
GLEAM에 입력한 모델

시뮬레이션 분석은 시간의 함수이기 때문에 바이러스 전염 시초일을 통일 하는 것이 다른 연구들과의 비교를 위해서도 중요하다. 홍콩과기대가 가장 최근에 올린 논문의 연구데이터 확보 최초 시작일인 2019년 12월 31일을 기점으로 해서 2020년 5월 까지의 신규 및 누적 감염자(확진자 아님)수를 분석하면 다음과 같다.

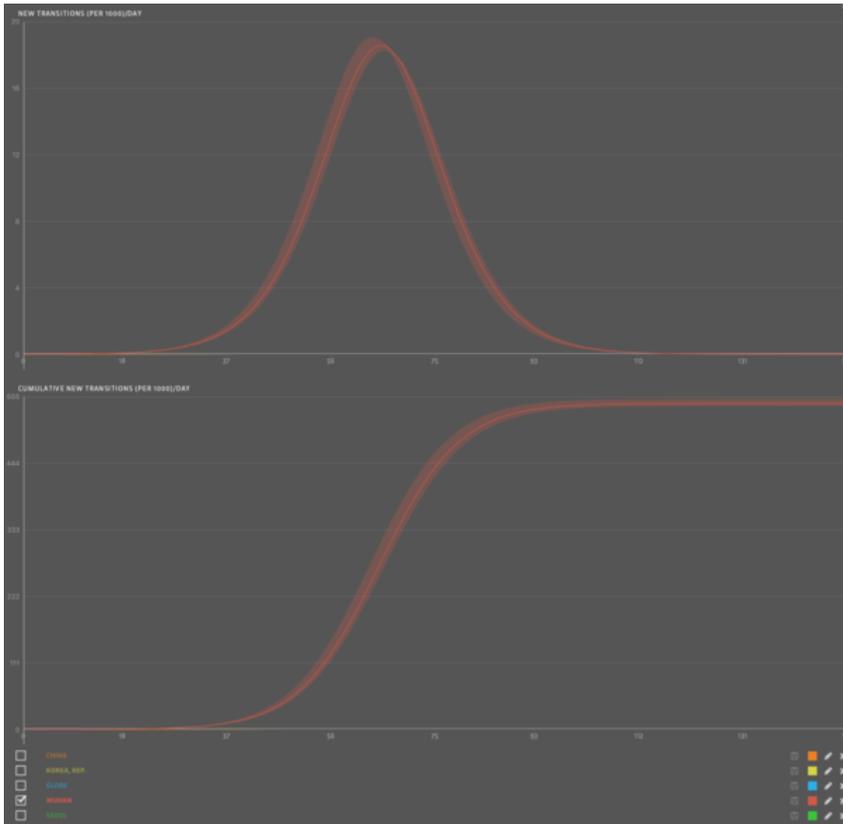


신종 코로나 바이러스 확산에 대한 GLEAM 분석 결과

예상 신규 감염자수 확산 순위는 중국→태국→인도네시아→말레이시아→베트남→싱가포르→캄보디아→필리핀 순이다. 한편 2020년 2월 2일 오후 3시 현재 존스홉킨스 대학의 신종 코로나 확진자수 통계로 보면 순위는 중국→일본→태국→싱가포르→한국→홍콩→호주→대만→말레이시아→베트남 순이다.

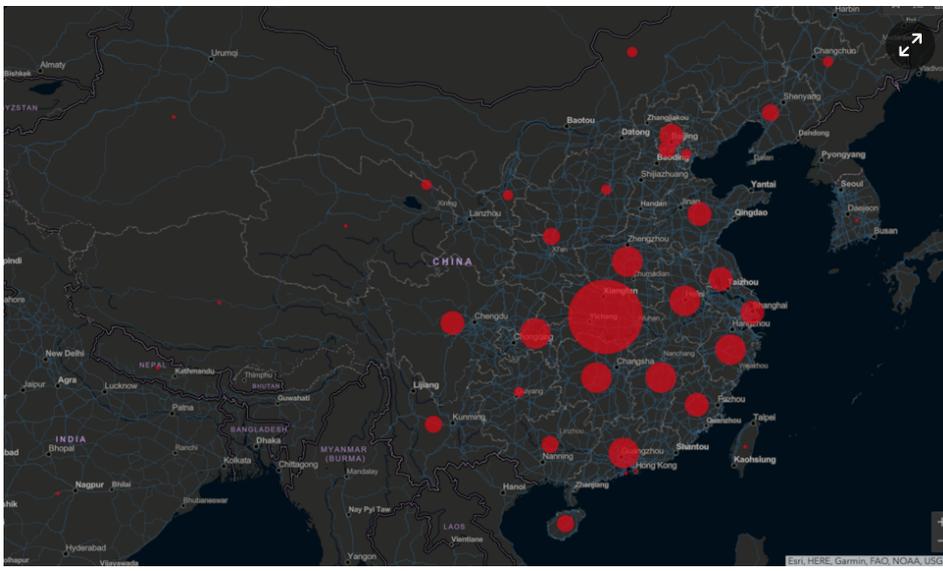


GLEAM 분석결과, 4월에는 중국내 신규 감염자수는 정점을 보일 것으로 분석 되었다. 파란색선은 글로벌 분석 추정치다.

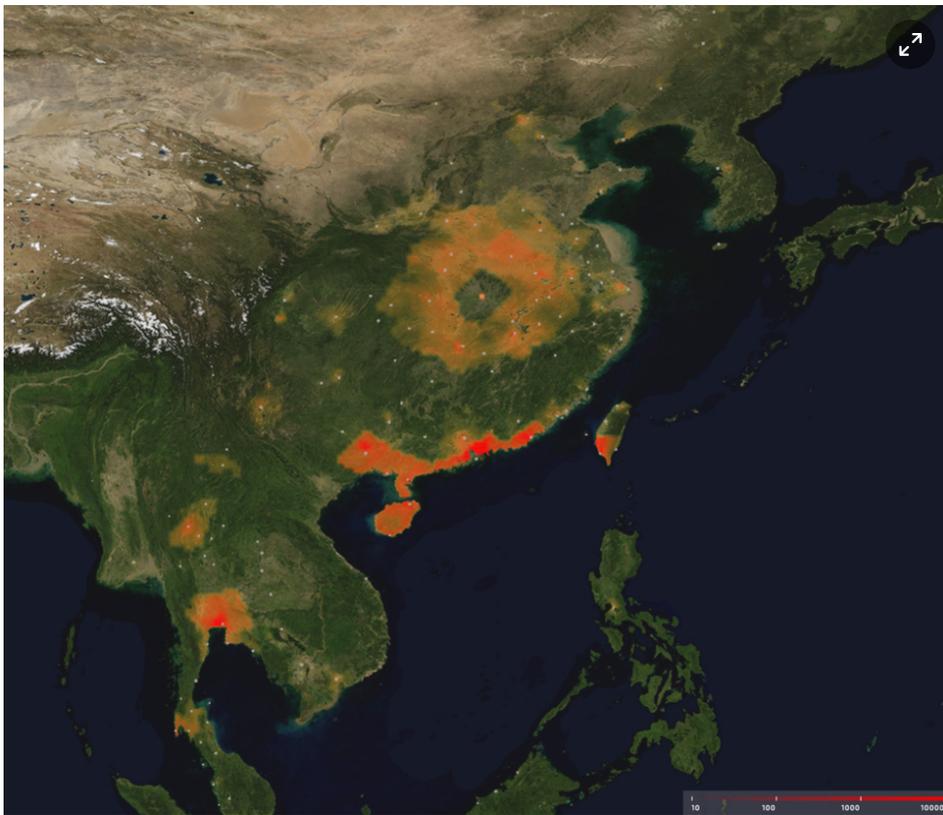


중국 우한시의 신규 및 누적 감염자수 분석이다. 우한의 경우는 2-3월이 신규 감염자수가 정점에 도달할 것으로 예상된다.

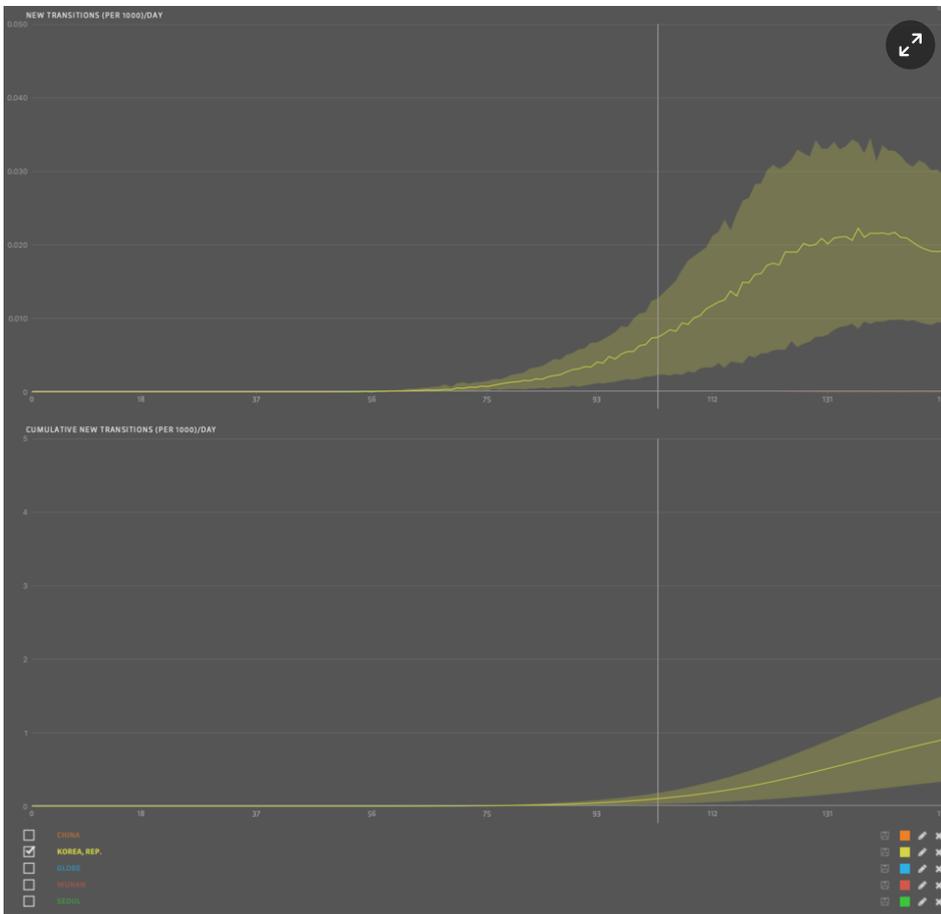
해당 그래프는 인구 1천명당 감염자수로서 2020년 2월 1일 현재 우한내에 약 7만5915명의 감염자가 있을 것으로 추정된 홍콩과기대의 논문데이터와 유사하다. 필자의 분석으로는 우한시의 인구를 천만명으로 가정할 경우 앞에서 가정한 시작일(2019년 12월 31일) 기준으로 약 **6만 7천명의 감염자가 있는 것으로 추정**되었다.



존스홉킨스대학에서 분석한 2020년 2월 2일 오후 3시 기준 확진자 지역 분포



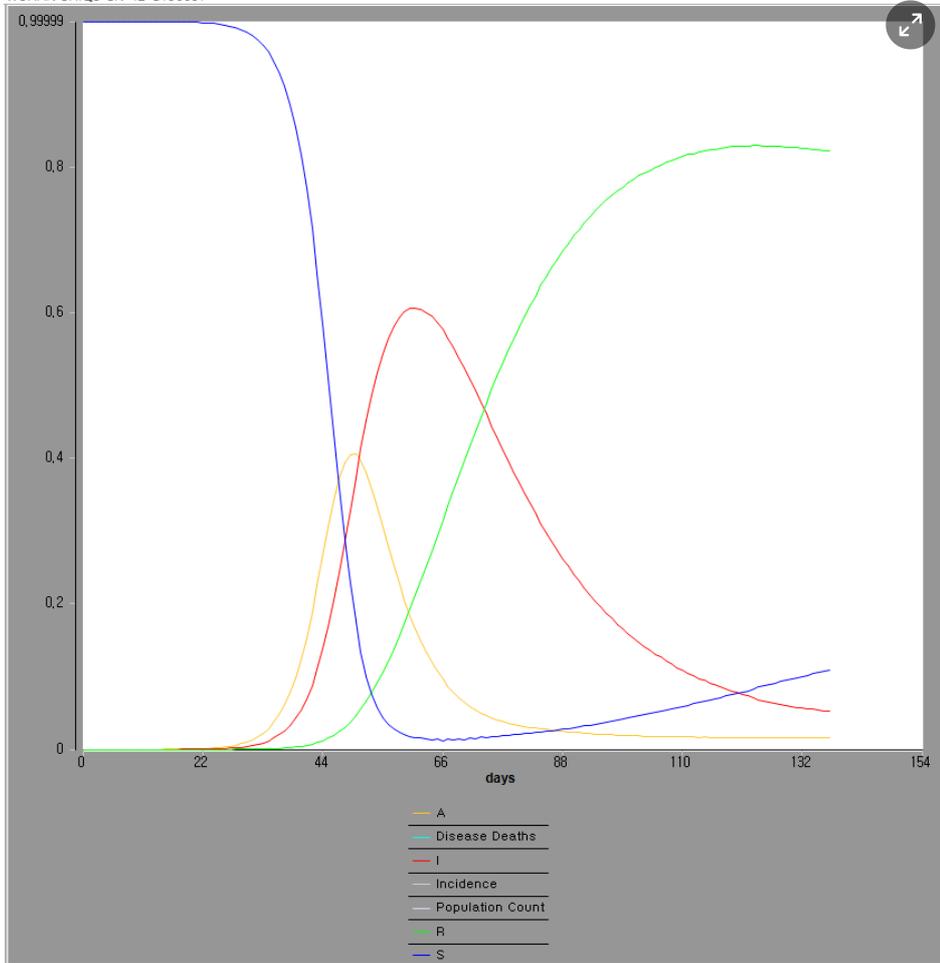
뉴스톱이 분석한 2020년 2월 2일 기준 감염자 확산 추정 분석. 한국 일본도 소규모 감염 상태로 분석되었다.



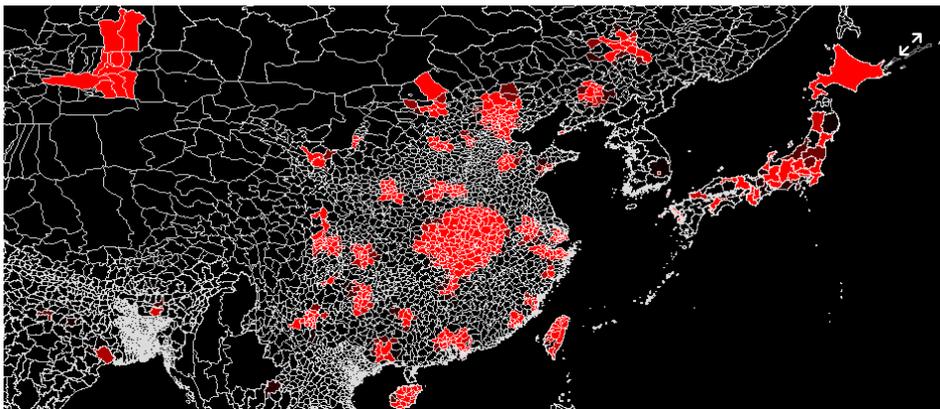
우리나라의 향후 5개월 후 누적 감염자수는 인구수 1천명당 평균 1명 그리고 신규 감염자수는 0.02명 수준으로 분석 되었다.

다음으로는 한국의 확산 추정이다. 서울 인구수를 1천만명으로 가정하면 약 1만명 정도 누적 감염자수까지 증가될 것으로 분석되었다. 단, 이 수치는 어디까지나 감염자 예측 수치로서 이들 중에는 증상이 나타나는 경우도 있고 모르고 지나가는 경우도 있고 자연회복 되는 경우도 있을 것이다. 또한 우리나라의 경우 전세계적으로도 공중보건 및 의료 수준이 최상위권이기 때문에 현재로서는 기존 독감 감염자 통계와 비교해도 위기상황은 분석 수치로도 예측되지는 않는다.

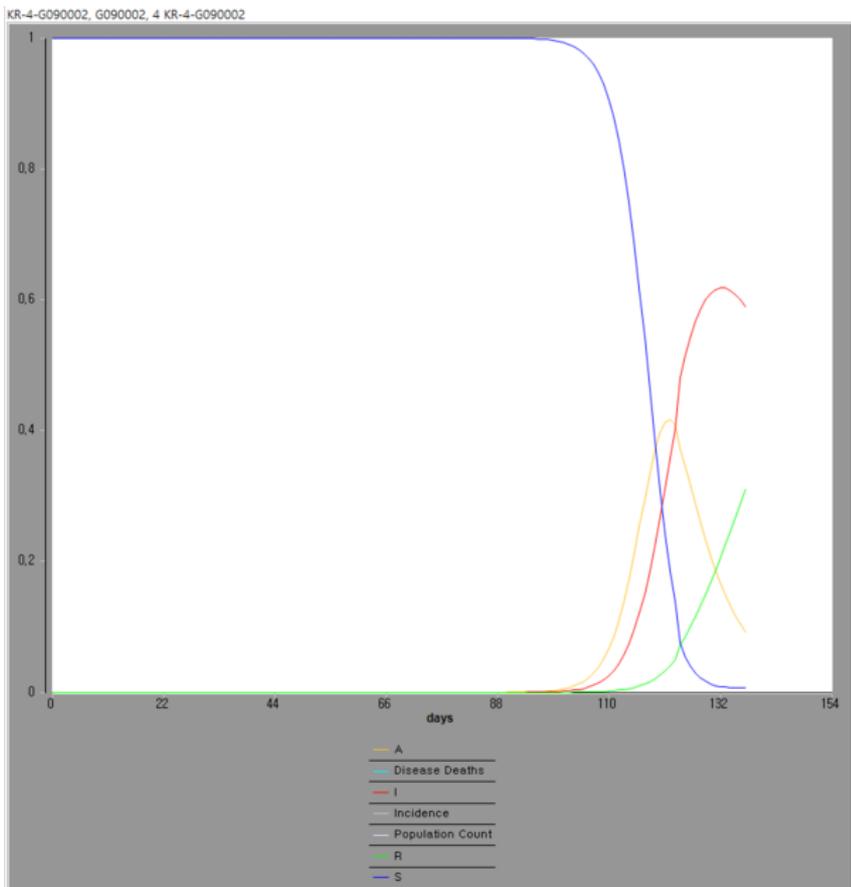
다음은 미국에서 주로 사용하는 STEM의 시뮬레이션 결과이다. 앞에서 홍콩과기대가 분석한 한 명의 감염자가 몇 명을 전염시킬 수 있는지 수치화한 기초재생산수(basic reproduction number)는 1인당 평균 2.68명으로 분석되었다. 뉴스톱에선 최근 추정치중 가장 높게 나왔던 3.4를 가지고 진행했다. 시작일은 신종 코로나 바이러스가 최초 보고된 것으로 알려진 2019년 12월 13일부터 분석했다. 시뮬레이션 결과치는 다소 과장된 것으로 보아도 무방하다.



앞의 GLEAM 결과와 비슷하게 우한시의 경우 2~3월이 신규 감염자수가 정점이 될 것으로 분석 되었다



STEM으로 돌려본 결과 2월 현재 우리나라보다 일본의 확산 가능성이 더 높게 나왔다. 참고로 2020년 2월 2일 현재 확진자수는 일본이 20명으로 중국 다음으로 2순위를 보여주고 있다.

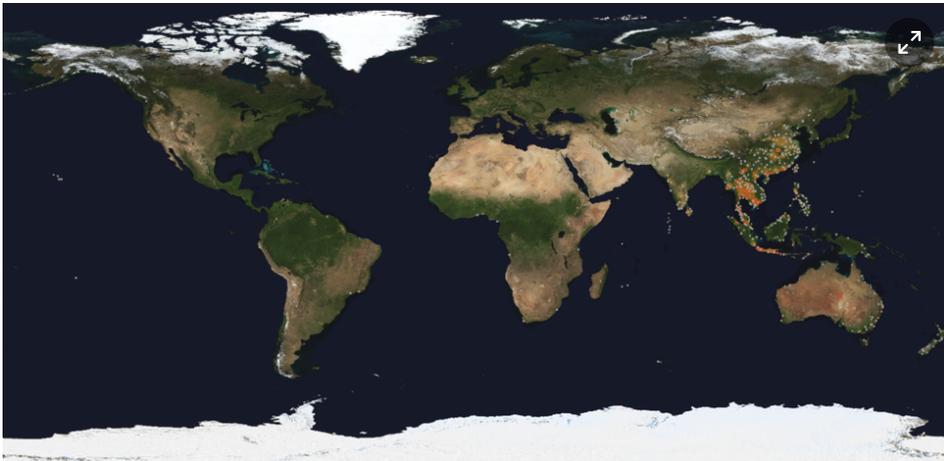


STEM으로 분석한 한국의 신규 감염자 수는 앞의 GLEAM 분석 결과와 유사하게 5월~6월에 정점을 보일 것으로 추정된다.

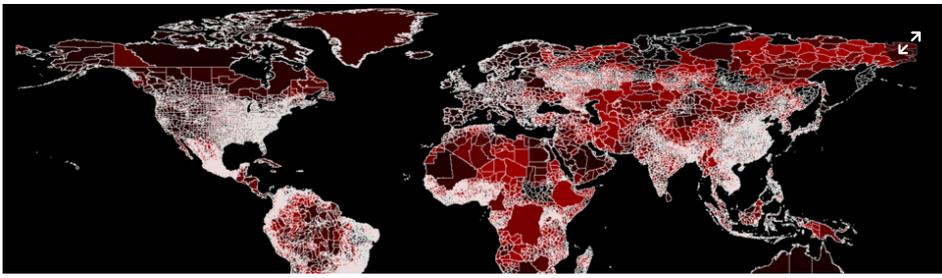
신종 코로나 바이러스의 전지구적 확산 시뮬레이션을 2020년 6월까지 진행하였다. GLEAM의 시뮬레이션 결과, 대부분의 감염자 확산은 중국과 그 인접국가였다.(GLEAM의 경우 경미하지만 무증상 감염 가능성에 대한 인자도 반영해서 시뮬레이션), 다행히 한국은 피해가 경미한 것으로는 예측되었다.

STEM의 분석결과는 좀더 광범위한 감염자 확산을 보여 주고 있다. 시간이 지날수록 미국 및 러시아와 아프리카 지역까지도 확산 위험에 처해질 수 있음을 예측하고 있다.(STEM의 경우 각국의 입국금지와 같은 봉쇄는 3월로 가정해서 시뮬레이션 한 결과)

많은 국가가 입국금지와 같은 선제적인 조치를 취하고 있어 신종 코로나 바이러스 감염증이 지역적 확산(Epidemic)이 될지 전지구적 확산(Pandemic)이 될지는 좀 더 지켜볼 필요가 있다.



GLEAM의 시뮬레이션 결과는 대부분의 감염자 확산은 중국과 그 인접국가였다. 다행히 우리나라는 피해가 경미한 것으로는 예측 되고 있다.



STEM의 분석결과는 좀더 광범위한 감염자 확산을 보여 주고 있다. 시간이 지날 수록 미국 및 러시아와 아프리카 지역까지도 확산 위험에 처해질 수 있음을 예측하고 있다. (붉은색의 농도는 구분을 명확하게 하기 위하여 GAIN을 올린 것으로 심각성으로 인지할 필요는 없다)

4. 분석결과 시사점 및 한계

복잡한 수학 알고리즘에서 출발한 감염병 확산 예측 기술은 각 나라의 인구통계, 교통 및 통근, 의료보건의수 등과 같은 빅데이터를 수용함으로써 좀더 현실적인 예측이 가능해지고 있는 추세이다. 감염병이 빠르게 전지구적으로 확산되고 다양한 변종 바이러스들이 창궐하는 시대에 빠른 정책적 판단이 가능하다는 점에서 이 기술의 의의가 있다. 다만 시뮬레이션은 수 많은 가정과 단순화의 결과물이기 때문에 정책결정에 있어서 전적으로 의존해서는 안된다. 국경봉쇄, 여행통제 등을 결정할 때는 좀 더 면밀한 검토와 전문가의 조언이 필요하다.

우한에서 발생한 신종 코로나 바이러스가 한국에 미치는 영향은 과거 사스나 메르스 사태때 보다는 제한적일 것으로 분석되었다. 중국 이외 지역에서의 확진 후 사망자 수가 아직은 유의미하지 않은 점도 충분한 역학관계를 해석하기에는 정보가 부족한 점도 있다. 특히 우리나라는 보건의료환경 인프라가 상위권 국가수준이기 때문에 정부를 믿고 개인위생을 철저히 하면 된다. 그러나 앞으로 2-3개월이 우리에게도 감염자 관리상의 중요한 변곡점이 될 것으로 분석되었다. 긴장을 늦출 수는 없는 상황이다.

아직까지는 중국이외의 지역에서도 급격한 사망자 확산이 일어나는 세계적 유행병(Pandemic)으로 보기는 무리가 있다. 각국이 과거 사스나 메르스 사태를 거치면서 충분한 학습을 하였기 때문에 비교적 과거보다 빠른 선제적 대응을 하고 있어 급격한 확산을 막고 있는 것으로 추정된다. 문제는 중국내에서의 감염속도와 사망률이 높다는 것과 의료보조 인프라가 충분하지 않은 저개발국가로의 확산 위험성이다. 중국은 2-3월이 감염자 확산의 정점을 보일 것으로 분석되었다. 그만큼 2-3월은 중국이외의 지역으로의 감염확산이 늘어 날 것으로 볼 수 밖에 없다. 한국의 경우 일본과 같은 중국 이외의 지역 감염자들을 통한 국내 유입 우회 가능성도 제기된다. 12번 확진자의 경우 일본에서 전염되어 온 중국인이었다. 국가간 상시 협력과 비상대응체계를 갖추는 것이 필요해진 상황이다.

해당 분석은 AI-머신러닝 기반 데이터 분석 전문 기업 링크브릭스의 도움을 받아 진행하였습니다.

지윤성 saxoji@newstof.com *최근글보기*

드론/자동차/카메라 등 대한민국 남자라면 좋아할 아이템들에 관심이 많은, 게임-소프트웨어-클라우드 서비스 관련 회사의 창업자 및 임원을 지내며 정보격차 없는 낭만적인 IT 세상을 꿈꾸고 있습니다.

Tag #STEM #GLEAM #전염병 확산 예측 #신종 코로나바이러스

저작권자 © 뉴스톱 무단전재 및 재배포 금지

지윤성 팩트체커

[다른기사 보기](#)



관련기사

NEWSTOF 바이러스 확산 시뮬레이션 해보니...한국은 상대적 '안전지대'



- 신종 코로나 바이러스의 '9가지 진실'
- 빌 게이츠 재단이 6500만명 사망 예측? 전 세계 허위정보 팩트체크
- [주간팩트체크] 코로나 바이러스 때문에 쓰러지는 사람 영상은 가짜?
- 신종 코로나 바이러스로 뿌리채 흔들리는 '시진핑 체제'
- 참기름도 마늘도 소용없다...WHO가 공개한 '신종 코로나'에 대한 10가지 미신

당신만안본뉴스

신종 코로나 바이러스의 '9가지 진실'

바이러스 확산 시뮬레이션 해보니...한국은 상대적 '안전지대'

1931년 우리만의 중국인 혐오, '130여명 학살'로 이어지다

"미국 독감으로 신종코로나 몰타기?"

[주간팩트체크] 중국인 입국금지하면 코로나바이러스 막을 수 있을까?

[팩트체크] KIST에서 이름 삭제된 '조만'은 조국 딸 맞나?

[팩트체크] 신종 코로나보다 '미국 독감'이 더 심각하다?

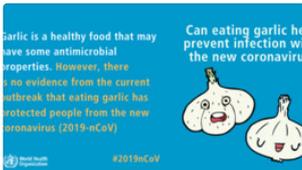


신종 코로나 바이러스의 '9가지 진실'

댓글 0 (뉴스톱 댓글달기는 회원으로 가입한 분만 가능합니다)

댓글입력

주요기사



참기름도 마늘도 소용없다...WHO가 공개한 '신종 코로



트랜스젠더 입학 반대? 생물학적으로 십여개의 성(Sex)이 존



[팩트체크] 신종 코로나보다 '미국 독감'이 더 심각하다?



"미국 독감으로 신종코로나 몰타기?"



미래한국당 대표 한선교... 與 "이적 권유 황교안 고발" 民



문재인 '코로나와 박근혜 메르스의 결정적 차이는 '0000'

서울특별시 중구 세종대로 16길 23 3층 뉴스톱 | 대표전화 : 070-7728-3300 | 팩스 : 02-2178-9130 | 청소년보호책임자 : 송영훈
 법인명 : (주)뉴스톱 | 제호 : 뉴스톱 | 등록번호 : 서울,04576 | 등록일 : 2017-06-15 | 발행일 : 2017-05-01 | 발행·편집인 : 김준일
 Copyright © 2020 뉴스톱. All rights reserved. mail to contact@newstof.com

NEWSTOF

매체소개 | 기사제보 | 광고문의 | 이용약관 | 개인정보처리방침 | 청소년보호정책 | 저작권보호정책 | 이메일무단수집거부

POWERED BY NE

